PROBABILITY INEQUALITIES AND TAIL ESTIMATES FOR METRIC SEMIGROUPS

APOORVA KHARE AND BALA RAJARATNAM
STANFORD UNIVERSITY

Abstract. We study probability inequalities leading to tail estimates in a general semigroup $\mathscr{G}$ with a translation-invariant metric $d_{\mathscr{G}}$. Using recent work [*Ann. Prob.*, to appear] that extends the Hoffmann-Jørgensen inequality to all metric semigroups, we obtain tail estimates and approximate bounds for sums of independent semigroup-valued random variables, their moments, and decreasing rearrangements. In particular, we obtain the "correct" universal constants in several cases, extending results in the Banach space literature by Johnson–Schechtman–Zinn [*Ann. Prob.* 13], Hitczenko [*Ann. Prob.* 22], and Hitczenko and Montgomery-Smith [*Ann. Prob.* 29]. Our results also hold more generally, in the minimal mathematical framework required to state them: metric semigroups $\mathscr{G}$. This includes all compact, discrete, or abelian Lie groups.

1. Introduction and main results

The goal of this paper is to extend the study of probability theory beyond the Banach space setting. In the present work, we estimate sums of independent random variables in several ways, under the most primitive mathematical assumptions required to state them. The setting is as follows.

Definition 1.1. A *metric semigroup* is defined to be a semigroup $(\mathscr{G}, \cdot)$ equipped with a metric $d_{\mathscr{G}} : \mathscr{G} \times \mathscr{G} \to [0, \infty)$ that is translation-invariant. In other words,

$$d_{\mathscr{G}}(ac, bc) = d_{\mathscr{G}}(a, b) = d_{\mathscr{G}}(ca, cb) \ \forall a, b, c \in \mathscr{G}.$$

(Equivalently, $(\mathscr{G}, d_{\mathscr{G}})$ is a metric space equipped with a associative binary operation such that $d_{\mathscr{G}}$ is translation-invariant.) Similarly, one defines a *metric monoid* and a *metric group*.

Metric groups are ubiquitous in probability theory, and subsume all compact and abelian Lie groups as well as normed linear spaces as special cases. More modern examples of recent interest are mentioned presently.

Date: October 7, 2016.
2010 *Mathematics Subject Classification.* Primary: 60E15, Secondary: 60B15, 60B10.
Key words and phrases. Metric semigroup, maximal inequality, Lévy inequality, Ottaviani–Skorohod inequality, Mogul'skii inequality, Lévy equivalence, tail estimate, moment estimate, decreasing rearrangement, universal constant.
A.K. and B.R. are partially supported by the following: US Air Force Office of Scientific Research grant award FA9550-13-1-0043, US National Science Foundation under grant DMS-0906392, DMS-CMG 1025465, AGS-1003823, DMS-1106642, DMS-CAREER-1352656, Defense Advanced Research Projects Agency DARPA YFA N66001-111-4131, the UPS Foundation, and SMC-DBNKY.

Now suppose $(\Omega, \mathscr{A}, \mu)$ is a probability space and $X_1, \dots, X_n \in L^0(\Omega, \mathscr{G})$ are $\mathscr{G}$-valued random variables. Fix $z_0, z_1 \in \mathscr{G}$ and define for $1 \leqslant j \leqslant n$:

$$S_j := X_1 X_2 \cdots X_j, \quad U_j := \max_{1 \leqslant i \leqslant j} d_{\mathscr{G}}(z_1, z_0 S_i), \quad Y_j := d_{\mathscr{G}}(z_0, z_0 X_j), \quad M_j := \max_{1 \leqslant i \leqslant j} Y_i. \tag{1.2}$$

In this paper we discuss bounds that govern the behavior of U_n – and consequently, of sums S_n of independent $\mathscr{G}$-valued random variables X_j – in terms of the variables X_j, and even Y_j or M_j. We are interested in a variety of bounds: (a) one-sided geometric tail estimates; (b) approximate two-sided bounds for tail probabilities; (c) approximate two-sided bounds for moments; and (d) comparison of moments. For instance, is it possible to obtain bounds for $\mathbb{E}_\mu[U_n^p]^{1/p}$ in terms of the tail distribution for U_n, or in terms of $\mathbb{E}_\mu[U_n^q]^{1/q}$ for $p, q > 0$? The latter question has been well-studied in the literature for Banach spaces, and universal bounds that grow at the "correct" rate have been obtained for all $q \gg 0$. We explore the question of obtaining correctly growing universal constants for metric semigroups, which include not only normed linear spaces and inner product spaces, but also all abelian and compact Lie groups. Our results show that the universal constants in such inequalities do not depend on the semigroup in question.

1.1. **Motivations.** Our motivations in developing probability theory in such general settings are both modern and classical. An increasing number of modern-day theoretical and applied settings require mathematical frameworks that go beyond Banach spaces. For instance, data and random variables may take values in manifolds such as (real or complex) Lie groups. Compact or abelian Lie groups also commonly feature in the literature, including permutation groups and other finite groups, lattices, orthogonal groups, and tori. In fact every abelian, Hausdorff, metrizable, topologically complete group G admits a translation-invariant metric [15], though this fails to hold for cancellative semigroups [16]. Certain classes of amenable groups are also metric groups (see [14] for more details). Other modern examples arise in the study of large networks and include the space of graphons with the cut norm, which arises naturally out of combinatorics and is related to many applications [19]. In a parallel vein, the space of labelled graphs $\mathscr{G}(V)$ on a fixed vertex set V is a 2-torsion metric group (see [11, 12]), hence does not embed into a normed linear space.

With these modern examples in mind, in this paper we develop novel techniques for proving maximal inequalities – as well as comparison results between tail distributions and various moments – for sums of independent random variables taking values in the aforementioned groups, which are not Banach spaces.

At the same time, we also have theoretical motivations in mind when developing probability theory on non-linear spaces such as $\mathscr{G}(V)$ and beyond. Throughout the past century, the emphasis in probability has shifted somewhat from proving results on stochastic convergence, to obtaining sharper and stronger bounds on random sums, in increasingly weaker settings. A celebrated achievement of probability theory has been to develop a rigorous and systematic framework for studying the behavior of sums of (independent) random variables; see e.g. [18]. In this vein, we provide *unifications* of our results on graph space with those in the Banach space literature, by proving them *in the most primitive mathematical framework possible.* In particular, our results apply to compact/abelian/discrete Lie groups, as well as normed linear spaces.

For example, maximal inequalities by Hoffmann-Jørgensen, Lévy, Ottaviani–Skorohod, and Mogul'skii require merely the notions of a metric and a binary associative operation to state them. Thus one only needs a separable metric semigroup $\mathscr{G}$ rather than a Banach space to state these inequalities. However, note that working in a metric semigroup raises technical questions. For instance, the lack of an identity element means one has to specify how to compute magnitudes of $\mathscr{G}$-valued random variables (before trying to bound or estimate them); also, it is not apparent how to define truncations of random variables. The lack of inverses, norms, or commutativity implies in particular that one cannot rescale or subtract random variables. Thus new methods need to be developed when the minimal mathematical structure of $\mathscr{G}$ makes it impossible to adopt and extend the existing proofs of the aforementioned results.

In the present work, we hope to show that the approach of working with arbitrary metric semigroups turns out to be richly rewarding in (i) obtaining the above (and other) results for non-Banach settings; (ii) unifying these results with the existing Banach space results in order to hold in the greatest possible generality; and (iii) further strengthening these unified versions where possible.

1.2. **Organization and results.** We now describe the organization and contributions of the present paper. In Section 2 we prove the Mogul'skii–Ottaviani–Skorohod inequalities for all metric semigroups $\mathscr{G}$. As an application, we show Lévy's equivalence for stochastic convergence in metric semigroups.

In Section 3, we come to our main goal in this paper, of estimating and comparing moments and tail probabilities for sums of independent $\mathscr{G}$-valued random variables. Our main tool is a variant of Hoffmann-Jørgensen's inequality for metric semigroups, which is shown in recent work [13]. The relevant part for our purposes is now stated.

Theorem 1.3 (Khare and Rajaratnam, [13])**.** *Notation as in Definition 1.1 and Equation* (1.2)*. Suppose $X_1, \dots, X_n \in L^0(\Omega, \mathscr{G})$ are independent. Fix integers $n_1, \dots, n_k \in \mathbb{N}$ and numbers $t_1, \dots, t_k, s \in [0, \infty)$, and define $I_0 := \{1 \leqslant i \leqslant k : \mathbb{P}_\mu(U_n \leqslant t_i)^{n_i - \delta_{i1}} \leqslant 1/n_i!\}$, where δ_{i1} denotes the Kronecker delta. Now if $\sum_{i=1}^k n_i \leqslant n+1$, then:*

$$\mathbb{P}_\mu\left(U_n > (2n_1 - 1)t_1 + 2\sum_{i=2}^k n_i t_i + \left(\sum_{i=1}^k n_i - 1\right) s\right)$$
$$\leqslant \mathbb{P}_\mu(M_n > s) + \mathbb{P}_\mu(U_n \leqslant t_1)^{\mathbf{1}_{1 \notin I_0}} \prod_{i \in I_0} \mathbb{P}_\mu(U_n > t_i)^{n_i} \prod_{i \notin I_0} \frac{1}{n_i!} \left(\frac{\mathbb{P}_\mu(U_n > t_i)}{\mathbb{P}_\mu(U_n \leqslant t_i)}\right)^{n_i}.$$

Remark that Theorem 1.3 generalizes the original Hoffmann-Jørgensen inequality in three ways: (i) mathematically it strengthens the state-of-the-art even for real variables; (ii) it unifies previous results by Johnson and Schechtman [*Ann. Prob.* 17], Klass and Nowicki [*Ann. Prob.* 28], and Hitczenko and Montgomery-Smith [*Ann. Prob.* 29] in the Banach space literature; and (iii) the result holds in the most primitive setting needed to state it, thereby being applicable also to e.g. Lie groups.

We now discuss several ways in which to estimate the size of sums of independent $\mathscr{G}$-valued random variables, for metric semigroups $\mathscr{G}$. We present two results in this section, corresponding to two of the estimation techniques discussed in the introduction. (For a third result, see Theorem 3.6.)

The **first** approach, informally speaking, uses the Hoffmann-Jørgensen inequality to generalize an upper bound for $\mathbb{E}_\mu[\|S_n\|^p]$ in terms of the quantiles of $\|S_n\|$ as well as $\mathbb{E}_\mu[M_n^p]$ – but now in the "minimal" framework of metric semigroups. More precisely, we show that controlling the behavior of X_n is equivalent to controlling S_n or U_n, for all metric semigroups.

Theorem A. *Suppose $A \subset \mathbb{N}$ is either $\mathbb{N}$ or $\{1,\dots,N\}$ for some $N \in \mathbb{N}$. Suppose $(\mathscr{G}, d_\mathscr{G})$ is a separable metric semigroup, $z_0, z_1 \in \mathscr{G}$, and $X_n \in L^0(\Omega, \mathscr{G})$ are independent for all $n \in A$. If $\sup_{n\in A} d_\mathscr{G}(z_1, z_0 S_n) < \infty$ almost surely, then for all $p \in (0,\infty)$,*

$$\mathbb{E}_\mu\left[\sup_{n\in A} d_\mathscr{G}(z_0, z_0X_n)^p\right] < \infty \qquad \Longleftrightarrow \qquad \mathbb{E}_\mu\left[\sup_{n\in A} d_\mathscr{G}(z_1, z_0S_n)^p\right] < \infty.$$

This result extends [7, Theorem 3.1] by Hoffmann-Jørgensen to the "minimal" framework of metric semigroups. The proofs of Theorem A and the next result use the notion of the quantile functions, or decreasing rearrangements, of $\mathscr{G}$-valued random variables:

Definition 1.4. Suppose $(\mathscr{G}, d_\mathscr{G})$ is a metric semigroup, and $X : (\Omega, \mathscr{A}, \mu) \to (\mathscr{G}, \mathscr{B}_\mathscr{G})$. We define the *decreasing* (or *non-increasing*) *rearrangement* of X to be the right-continuous inverse X^* of the function $t \mapsto \mathbb{P}_\mu\left(d_\mathscr{G}(z_0, z_0X) > t\right)$, for any $z_0 \in \mathscr{G}$. In other words, X^* is the real-valued random variable defined on $[0,1]$ with the Lebesgue measure, as follows:

$$X^*(t) := \sup\{y \in [0,\infty) : \mathbb{P}_\mu\left(d_\mathscr{G}(z_0, z_0X) > y\right) > t\}.$$

Note that X^* has exactly the same law as $d_\mathscr{G}(z_0, z_0X)$. Moreover, if $(\mathscr{G}, \|\cdot\|)$ is a normed linear space, then $d_\mathscr{G}(z_0, z_0X)$ can be replaced by $\|X\|$, and often papers in the literature refer to X^* as the decreasing rearrangement of $\|X\|$ instead of X itself. The convention that we adopt above is slightly weaker.

The **second** approach provides another estimate on the size of S_n through its moments, by comparing $\|S_n\|_q$ to $\|S_n\|_p$ – or more precisely, $\mathbb{E}_\mu[U_n^q]^{1/q}$ to $\mathbb{E}_\mu[U_n^p]^{1/p}$ – for $0 < p \leqslant q$. Moreover, the constants of comparison are universal, valid for all abelian semigroups and all finite sequences of independent random variables, and depend only on a threshold:

Theorem B. *Given $p_0 > 0$, there exist universal constants $c = c(p_0), c' = c'(p_0) > 0$ depending only on p_0, such that for all choices of (a) separable abelian metric semigroups $(\mathscr{G}, d_\mathscr{G})$, (b) finite sequences of independent $\mathscr{G}$-valued random variables $X_1,\dots,X_n$, (c) $q \geqslant p \geqslant p_0$, and (d) $\epsilon \in (-q, \log(16)]$, we have*

$$\begin{aligned}\mathbb{E}_\mu[U_n^q]^{1/q} &\leqslant c\frac{q}{\max(p, \log(\epsilon+q))}(\mathbb{E}_\mu[U_n^p]^{1/p} + M_n^*(e^{-q}/8)) + c\mathbb{E}_\mu[M_n^q]^{1/q}\\ &\leqslant c'\frac{q}{\max(p, \log(\epsilon+q))}(\mathbb{E}_\mu[U_n^p]^{1/p} + \mathbb{E}_\mu[M_n^q]^{1/q}) \qquad \textit{if } \epsilon \geqslant \min(1, e-p_0).\end{aligned}$$

Moreover, we may choose

$$c'(p_0) = c(p_0) \cdot \left(8^{1/p_0}e + \max(1, \frac{\log(\epsilon+p_0)}{p_0})\right).$$

Theorem B extends a host of results in the Banach space literature, including by Johnson–Schechtman–Zinn [*Ann. Prob.* 13], Hitczenko [*Ann. Prob.* 22], and Hitczenko and Montgomery-Smith [*Ann. Prob.* 29]. (See also [18, Theorem 6.20] and [17, Proposition 1.4.2].) Theorem B also yields the correct order of the constants as $q \to \infty$, as discussed by Johnson *et al* in *loc. cit.* where they extend previous work on Khinchin's

inequality by Rosenthal [22]. Moreover, all of these results are shown for Banach spaces. Theorem B holds additionally for all compact Lie groups, finite abelian groups and lattices, and spaces of labelled and unlabelled graphs.

2. Lévy's equivalence in metric semigroups

In this section we prove:

Theorem 2.1 (Lévy's Equivalence)**.** *Suppose $(\mathscr{G}, d_{\mathscr{G}})$ is a complete separable metric semigroup, $X_n : (\Omega, \mathscr{A}, \mu) \to (\mathscr{G}, \mathscr{B}_{\mathscr{G}})$ are independent, $X \in L^0(\Omega, \mathscr{G})$, and S_n is defined as in* (1.2)*. Then*

$$S_n \longrightarrow X \ a.s.\ P_\mu \qquad \Longleftrightarrow \qquad S_n \xrightarrow{P} X.$$

Moreover, if these conditions fail, then S_n diverges almost surely.

Special cases of this result have been shown in the literature. For instance, [2, §9.7] considers $\mathscr{G} = \mathbb{R}^n$. The more general case of a separable Banach space $\mathbb{B}$ was shown by Itô–Nisio [9, Theorem 3.1], as well as by Hoffmann-Jørgensen and Pisier [8, Lemma 1.2]. The most general version in the literature to date is by Tortrat, who proved the result for a complete separable metric group in [23]. Thus Theorem 2.1 is the closest to assuming only the minimal structure necessary to state the result (as well as to prove it).

In order to prove Theorem 2.1, we first study basic properties of metric semigroups. Note that for a metric group, the following is standard; see [15], for instance.

Lemma 2.2. *If $(\mathscr{G}, d_{\mathscr{G}})$ is a metric (semi)group, then the translation-invariance of $d_{\mathscr{G}}$ implies the "triangle inequality":*

$$d_{\mathscr{G}}(y_1 y_2, z_1 z_2) \leqslant d_{\mathscr{G}}(y_1, z_1) + d_{\mathscr{G}}(y_2, z_2) \ \forall y_i, z_i \in \mathscr{G}, \tag{2.3}$$

and in turn, this implies that each (semi)group operation is continuous.

If instead $\mathscr{G}$ is a group equipped with a metric $d_{\mathscr{G}}$, then except for the last two statements, any two of the following assertions imply the other two:

1. *$d_{\mathscr{G}}$ is left-translation invariant: $d_{\mathscr{G}}(ca, cb) = d_{\mathscr{G}}(a, b)$ for all $a, b, c \in \mathscr{G}$. In other words, left-multiplication by any $c \in \mathscr{G}$ is an isometry.*
2. *$d_{\mathscr{G}}$ is right-translation invariant.*
3. *The inverse map : $\mathscr{G} \to \mathscr{G}$ is an isometry. Equivalently, the triangle inequality* (2.3) *holds.*
4. *$d_{\mathscr{G}}$ is invariant under all inner/conjugation automorphisms.*

In order to show Theorem 2.1 for metric semigroups, we collect in the following proposition a few preliminary results from [14], and will use these below without further reference.

Proposition 2.4 ([14])**.** *Suppose $(\mathscr{G}, d_{\mathscr{G}})$ is a metric semigroup, and $a, b \in \mathscr{G}$. Then*

$$d_{\mathscr{G}}(a, ba) = d_{\mathscr{G}}(b, b^2) = d_{\mathscr{G}}(a, ab) \tag{2.5}$$

is independent of $a \in \mathscr{G}$. Moreover, a set $\mathscr{G}$ is a metric semigroup only if $\mathscr{G}$ is a metric monoid, or the set of non-identity elements in a metric monoid $\mathscr{G}'$. This is if and only if the number of idempotents in $\mathscr{G}$ is one or zero, respectively. Furthermore, the metric monoid $\mathscr{G}'$ is (up to a monoid isomorphism) the unique smallest element in the class of metric monoids containing $\mathscr{G}$ as a sub-semigroup.

Remark 2.6. If needed below, we will denote the unique metric monoid containing a given metric semigroup $\mathscr{G}$ by $\mathscr{G}' := \mathscr{G} \cup \{1'\}$. Note that the idempotent $1'$ may already be in $\mathscr{G}$, in which case $\mathscr{G} = \mathscr{G}'$. One consequence of Proposition 2.4 is that instead of working with metric semigroups, one can use the associated monoid $\mathscr{G}'$ instead. (In other words, the (non)existence of the identity is not an issue in many such cases.) This helps simplify other calculations. For instance, what would be a lengthy, inductive (yet straightforward) computation now becomes much simpler: for nonnegative integers k, l, and $z_0, z_1, \dots, z_{k+l} \in \mathscr{G}$, the triangle inequality (2.3) implies:

$$d_{\mathscr{G}}(z_0 \cdots z_k, z_0 \cdots z_{k+l}) = d_{\mathscr{G}'}(1', \prod_{i=1}^{l} z_{k+i}) \leqslant \sum_{i=1}^{l} d_{\mathscr{G}'}(1', z_{k+i}) = \sum_{i=1}^{l} d_{\mathscr{G}}(z_0, z_0 z_{k+i}).$$

2.1. **The Mogul'skii inequalities and proof of Lévy's equivalence.** Like Lévy's Equivalence (Theorem 2.1) and the Hoffmann-Jørgensen inequality (Theorem 1.3), many other maximal and minimal inequalities can be formulated using only the notions of a distance function and of a semigroup operation. We now extend to metric semigroups two inequalities by Mogul'skii, which were used in [20] to prove a law of the iterated logarithm in normed linear spaces. The following result will be useful in proving Theorem 2.1.

Proposition 2.7 (Mogul'skii–Ottaviani–Skorohod inequalities)**.** *Suppose $(\mathscr{G}, d_{\mathscr{G}})$ is a separable metric semigroup, $z_0, z_1 \in \mathscr{G}$, $a, b \in [0, \infty)$, and $X_1, \dots, X_n \in L^0(\Omega, \mathscr{G})$ are independent. Then for all integers $1 \leqslant m \leqslant n$,*

$$\mathbb{P}_\mu \left(\min_{m \leqslant k \leqslant n} d_{\mathscr{G}}(z_1, z_0 S_k) \leqslant a \right) \cdot \min_{m \leqslant k \leqslant n} \mathbb{P}_\mu \left(d_{\mathscr{G}}(S_k, S_n) \leqslant b \right) \leqslant \mathbb{P}_\mu \left(d_{\mathscr{G}}(z_1, z_0 S_n) \leqslant a + b \right),$$
$$\mathbb{P}_\mu \left(\max_{m \leqslant k \leqslant n} d_{\mathscr{G}}(z_1, z_0 S_k) \geqslant a \right) \cdot \min_{m \leqslant k \leqslant n} \mathbb{P}_\mu \left(d_{\mathscr{G}}(S_k, S_n) \leqslant b \right) \leqslant \mathbb{P}_\mu \left(d_{\mathscr{G}}(z_1, z_0 S_n) \geqslant a - b \right).$$

These inequalities strengthen [20, Lemma 1] from normed linear spaces to arbitrary metric semigroups. Also note that the second inequality generalizes the *Ottaviani–Skorohod inequality* to all metric semigroups. Indeed, sources such as [2, §9.7.2] prove this result in the special case $\mathscr{G} = (\mathbb{R}^n, +), z_0 = z_1 = 0, m = 1, a = \alpha + \beta, b = \beta$, with $\alpha, \beta > 0$.

We omit the proof of Proposition 2.7 for brevity as it involves standard arguments. Using this result, one can now prove Theorem 2.1. The idea is to use the approach in [2]; however, it needs to be suitably modified in order to work in the current level of generality.

Proof of Theorem 2.1. The forward implication is true in much greater generality. Conversely, we claim that S_i is Cauchy almost everywhere, if it converges in probability to X. Given $\epsilon, \eta > 0$, the assumption and definitions imply that there exists $n_0 \in \mathbb{N}$ such that

$$\mathbb{P}_\mu \left(d_{\mathscr{G}}(S_m, X) \geqslant \epsilon/8 \right) < \frac{\eta}{2(1+\eta)}, \quad \forall m \geqslant n_0.$$

This implies: $\mathbb{P}_\mu \left(d_{\mathscr{G}}(S_m, S_n) \geqslant \epsilon/4 \right) < \dfrac{\eta}{1+\eta}$ $\forall n \geqslant m \geqslant n_0$. Now define $S'_i := \prod_{j=1}^{i} X_{n_0+j}$. Fix $n > n_0$ and apply Proposition 2.7 to $\{X_{n_0+i} : i \in \mathbb{N}\}$ with $m = 1, a = \epsilon/2, b = \epsilon/4$,

and $z_0 = z_1$:

$$\begin{aligned}&\mathbb{P}_\mu\left(\max_{n_0+1\leqslant m\leqslant n} d_{\mathscr{G}}(S_{n_0}, S_m) \geqslant \epsilon/2\right) = \mathbb{P}_\mu\left(\max_{1\leqslant i\leqslant n-n_0} d_{\mathscr{G}'}(z_0, z_0 S'_i) \geqslant \epsilon/2\right)\\ \leqslant\ & \frac{\mathbb{P}_\mu\left(d_{\mathscr{G}'}(z_0, z_0 S'_{n-n_0}) \geqslant \epsilon/4\right)}{1 - \max_{1\leqslant i\leqslant n-n_0} \mathbb{P}_\mu\left(d_{\mathscr{G}'}(S'_i, S'_{n-n_0}) \geqslant \epsilon/4\right)} < \frac{\eta/(1+\eta)}{1-\eta/(1+\eta)} = \eta.\end{aligned}$$

Now define $Q_{n_0} := \sup_{n>n_0} d_{\mathscr{G}}(S_{n_0}, S_n)$ and $\delta_{n_0} := \sup_{n>m>n_0} d_{\mathscr{G}}(S_m, S_n)$. Then $\delta_{n_0} \leqslant 2Q_{n_0}$; moreover, taking the limit of the above inequality as $n \to \infty$ yields:

$$\mathbb{P}_\mu\left(Q_{n_0} \geqslant \epsilon/2\right) \leqslant \eta \quad \implies \quad \mathbb{P}_\mu\left(\delta_{n_0} \geqslant \epsilon\right) \leqslant \eta.$$

But then $\mathbb{P}_\mu\left(\sup_{n>m} d_{\mathscr{G}}(S_m, S_n) \geqslant \epsilon\right) \leqslant \eta$ for all $m > n_0$. Thus, S_n is Cauchy almost everywhere. Since $\mathscr{G}$ is complete, the result now follows from [2, Lemma 9.2.4]; that the almost sure limit is X is because $S_n \xrightarrow{P} X$. Finally, S_n either converges almost surely or diverges almost surely by the Kolmogorov 0-1 law, which concludes the proof. □

We remark for completeness that the other Lévy equivalence has been addressed in [1, 3, 23] for various classes of topological groups. See also [21] for a variant in discrete completely simple semigroups, [2, 9] for Banach space versions, and [14] for a version over any normed abelian metric group.

3. Measuring the magnitude of sums of independent random variables

We now prove Theorems A and B using the Hoffmann-Jørgensen inequality in Theorem 1.3. Recall that the Banach space version of this inequality is extremely important in the literature and is widely used in bounding sums of independent Banach space-valued random variables. Having proved Theorem 1.3, an immediate application of our main result is in obtaining the first such bounds for metric semigroups $\mathscr{G}$. We also provide uniformly good L^p-bounds and tail probability bounds on sums S_n of independent $\mathscr{G}$-valued random variables.

3.1. **An upper bound by Hoffmann-Jørgensen.**

In this subsection we prove Theorem A. The proof uses basic properties of decreasing rearrangements (see Definition 1.4), which we record here and use below, possibly without reference.

Proposition 3.1. *Suppose $X, Y : (\Omega, \mathscr{A}, \mu) \to [0, \infty)$ are random variables, and*

$$x, \alpha, \beta, \gamma > 0, \quad t \in [0, 1].$$

1. *$X^*(t) \leqslant x$ if and only if $\mathbb{P}_\mu\left(X > x\right) \leqslant t$.*
2. *$X^*(t)$ is decreasing in $t \in [0, 1]$ and increasing in $X \geqslant 0$.*
3. *$(X/x)^*(t) = X^*(t)/x$.*
4. *Suppose $\mathbb{P}_\mu\left(X > x\right) \leqslant \beta \mathbb{P}_\mu\left(Y > \gamma x\right)$ for all $x > 0$. Then for all $p \in (0, \infty)$ and $t \in (0, 1)$,*
$$\mathbb{E}_\mu[Y^p] \geqslant \beta^{-1}\gamma^p \mathbb{E}_\mu[X^p], \qquad \mathbb{E}_\mu[X^p] \geqslant t X^*(t)^p.$$
5. *Fix finitely many tuples of positive constants $(\alpha_i, \beta_i, \gamma_i, \delta_i)_{i=1}^N$, and real-valued non-decreasing functions f_i such that for all $x > 0$ there exists at least one i such that*
$$f_i(\mathbb{P}_\mu\left(X > \alpha_i x\right)) \leqslant \beta_i \mathbb{P}_\mu\left(Y > \gamma_i x\right)^{\delta_i}. \tag{3.2}$$

Then

$$X^*(t) \leqslant \max_{1 \leqslant i \leqslant N} \frac{\alpha_i}{\gamma_i} Y^*((f_i(t)/\beta_i)^{1/\delta_i}). \tag{3.3}$$

If on the other hand (3.2) *holds for all* i*, then* $X^*(t) \leqslant \min_{1 \leqslant i \leqslant N} \frac{\alpha_i}{\gamma_i} Y^*((f_i(t)/\beta_i)^{1/\delta_i})$.

Using these arguments, we now sketch the proof of one of the main results in this paper.

Proof of Theorem A. The backward implication is easy. Conversely, first claim that controlling sums of $\mathscr{G}$-valued L^p random variables in probability (i.e., in L^0) allows us to control these sums in L^p as well, for $p > 0$. Namely, we make the following claim:

Suppose $(\mathscr{G}, d_{\mathscr{G}})$ *is a separable metric semigroup,* $p \in (0, \infty)$*, and* $X_1, \dots, X_n \in L^p(\Omega, \mathscr{G})$ *are independent. Now fix* $z_0, z_1 \in \mathscr{G}$ *and let* S_k, U_n, M_n *be as in Definition 1.1 and Equation* (1.2)*. Then,*

$$\mathbb{E}_\mu[U_n^p] \leqslant 2^{1+2p}(\mathbb{E}_\mu[M_n^p] + U_n^*(2^{-1-2p})^p).$$

Note that the claim is akin to the upper bound by Hoffmann-Jørgensen that bounds $\mathbb{E}_\mu[\|S_n\|^p]$ in terms of $\mathbb{E}_\mu[M_n^p]$ and the quantiles of $\|S_n\|$ for Banach space-valued random variables. (See [7, proof of Theorem 3.1] and [4, Lemma 3.1].) We omit its proof for brevity, as a similar statement is asserted in [18, Proposition 6.8]. Now given the claim, define:

$$\begin{aligned} t_n :=& U_n^*(2^{-1-2p}) \quad (n \in A), \qquad & U_A := \sup_{n \in A} d_{\mathscr{G}}(z_1, z_0 S_n), \\ M_A :=& \sup_{n \in A} d_{\mathscr{G}}(z_0, z_0 X_n), & t_A := U_A^*(2^{-1-2p}), \end{aligned} \tag{3.4}$$

as above, where we also use the assumption that $U_A < \infty$ almost surely. Now for all $n \in A$, compute using the above claim and elementary properties of decreasing rearrangements:

$$\mathbb{E}_\mu[U_n^p] \leqslant 2^{1+2p}\mathbb{E}_\mu[M_n^p] + 2(4t_n)^p \leqslant 2^{1+2p}\mathbb{E}_\mu[M_A^p] + 2(4t_A)^p.$$

This concludes the proof if A is finite; for $A = \mathbb{N}$, use the Monotone Convergence Theorem for the increasing sequence $0 \leqslant U_n^p \to U_A^p$. ☐

3.2. **Two-sided bounds and L^p norms.** We now formulate and prove additional results that control tail behavior for metric semigroups and monoids – specifically, M_A, U_n, U_n^*. This includes proving our other main result, Theorem B. We begin by setting notation.

Definition 3.5. Suppose $\mathscr{G}$ is a metric semigroup.

(1) Given $X_n \in L^0(\Omega, \mathscr{G})$ as above, for all n in a finite or countable set A, define the random variable $\ell_X = \ell_{(X_n)} : [0,1] \to [0, \infty]$ via:

$$\ell_X(t) := \inf\{y > 0 \ : \ \sum_{n \in A} \mathbb{P}_\mu \left(d_{\mathscr{G}}(z_0, z_0 X_n) > y\right) \leqslant t\}.$$

As indicated in [6, §2], one then has: $\mathbb{P}_\mu(\ell_X > x) = \sum_{n \in A} \mathbb{P}_\mu \left(d_{\mathscr{G}}(z_0, z_0 X_n) > x\right)$.

(2) Two families of variables $P(t)$ and $Q(t)$ are said to be *comparable*, denoted by $P(t) \approx Q(t)$, if there exist constants $c_1, c_2 > 0$ such that $c_1^{-1} P(t) \leqslant Q(t) \leqslant c_2 P(t)$. The c_i are called the "constants of approximation".

(3) (For the remaining definitions, assume $(\mathscr{G}, 1_{\mathscr{G}}, d_{\mathscr{G}})$ is a separable metric monoid.) Given $t \geqslant 0$ and a random variable $X \in L^0(\Omega, \mathscr{G})$, define its *truncation* to be:

$$X(t) := \begin{cases} 1_{\mathscr{G}}, & \text{if } d_{\mathscr{G}}(1_{\mathscr{G}}, X) > t, \\ X, & \text{otherwise.} \end{cases}$$

(4) Given variables $X_1, \dots, X_n : \Omega \to \mathscr{G}$, and $r \in (0,1)$, define:

$$U_n'(r) := \max_{1 \leqslant k \leqslant n} d_{\mathscr{G}}(1_{\mathscr{G}}, \prod_{i=1}^k X_i(\ell_X(r))).$$

The following estimate on tail behavior compares U_n with its decreasing rearrangement.

Theorem 3.6. *Given $p_0 > 0$, there exist universal constants of approximation (depending only on p_0), such that for all $p \geqslant p_0$, separable abelian metric monoids $(\mathscr{G}, 1_{\mathscr{G}}, d_{\mathscr{G}})$, and finite sequences $X_1, \dots, X_n$ of independent $\mathscr{G}$-valued random variables (for any $n \in \mathbb{N}$),*

$$\mathbb{E}_\mu[U_n^p]^{1/p} \approx U_n^*(e^{-p}/4) + \mathbb{E}_\mu[\ell_X^p]^{1/p} \approx (U_n'(e^{-p}/8))^*(e^{-p}/4) + \mathbb{E}_\mu[\ell_X^p]^{1/p},$$

where U_n, U_n' were defined in Equation (1.2) and Definition 3.5 respectively.

For real-valued X, the expression $\mathbb{E}[|X|^p]^{1/p}$ is also denoted by $\|X\|_p$ in the literature.

To show Theorem 3.6, we require some preliminary results which provide additional estimates to govern tail behavior, and which we now collect before proving the theorem. As these preliminaries are often extensions to metric semigroups of results in the Banach space literature, we sketch or omit their proofs when convenient.

The first result obtains two-sided bounds to control the behavior of the "maximum magnitude" M_A (cf. Equation (3.4)).

Proposition 3.7. *Suppose $\{X_n : n \in A\}$ is a (finite or countably infinite) sequence of independent random variables with values in a separable metric semigroup $(\mathscr{G}, d_{\mathscr{G}})$.*

(1) *For all $t \in (0,1)$, $\ell_X(2t) \leqslant \ell_X(t/(1-t)) \leqslant M_A^*(t) \leqslant \ell_X(t)$.*
(2) *Suppose $X_n \in L^p(\Omega, \mathscr{G})$ for some $p > 0$ (and for all $n \in A$). For all $t > 0$, define:*

$$\Psi_X(t) := p \sum_{n \in A} \int_{\ell_X(t)}^\infty u^{p-1} \mathbb{P}_\mu \left(d_{\mathscr{G}}(z_0, z_0 X_n) > u \right) \, du.$$

Then, $$\frac{t\ell_X(t)^p + \Psi_X(t)}{1+t} \leqslant \mathbb{E}_\mu[M_A^p] \leqslant \ell_X(t)^p + \Psi_X(t).$$

Proof. The first part follows [6, Proposition 1] (using a special case of Equation (3.3)). For the second, follow [4, Lemma 3.2]; see also [18, Lemma 6.9]. □

We next discuss a consequence of Hoffmann-Jørgensen's inequality for metric semigroups, Theorem 1.3, which can be used to bound the L^p-norms of the variables U_n – or more precisely, to relate these L^p-norms to the tail distributions of U_n via U_n^*.

Lemma 3.8. *(Notation as in Definition 1.1 and Equation (1.2).) There exists a universal positive constant c_1 such that for any $0 \leqslant t \leqslant s \leqslant 1/2$, any separable metric semigroup $(\mathscr{G}, d_{\mathscr{G}})$ with elements z_0, z_1, and any sequence of independent $\mathscr{G}$-valued random variables $X_1, \dots, X_n$,*

$$U_n^*(t) \leqslant c_1 \frac{\log(1/t)}{\max\{\log(1/s), \log\log(4/t)\}} (U_n^*(s) + M_n^*(t/2)).$$

Proof. We begin by writing down a consequence of Theorem 1.3:

$$\mathbb{P}_\mu\left(U_n > (3K-1)t\right) \leqslant \frac{1}{K!}\left(\frac{\mathbb{P}_\mu\left(U_n > t\right)}{\mathbb{P}_\mu\left(U_n \leqslant t\right)}\right)^K + \mathbb{P}_\mu\left(M_n > t\right), \quad \forall t > 0,\ \forall K, n \in \mathbb{N}. \quad (3.9)$$

If $\mathbb{P}_\mu\left(U_n > t\right) \leqslant 1/2$, then this quantity is further dominated by

$$2\max\left\{\mathbb{P}_\mu\left(M_n > t\right), \frac{1}{K!}(2\mathbb{P}_\mu\left(U_n > t\right))^K\right\}.$$

Now carry out the steps mentioned in the proof of [6, Corollary 1]. □

The final preliminary result is proved by adapting the proofs of [6, Lemma 3 and Corollary 2] to metric monoids.

Proposition 3.10. *Suppose $(\mathscr{G}, 1_\mathscr{G}, d_\mathscr{G})$ is a separable metric monoid and $X_1, \dots, X_n : \Omega \to \mathscr{G}$ is a finite sequence of independent $\mathscr{G}$-valued random variables. For $r \in (0,1)$, define:*

$$U_n''(r) := \max_{1 \leqslant k \leqslant n} d_\mathscr{G}(1_\mathscr{G}, \prod_{i=1}^k X_i'(\ell_X(r))),$$

where $X_i'(t)$ equals $1_\mathscr{G}$ if $d_\mathscr{G}(1_\mathscr{G}, X_i) \leqslant t$, and X_i otherwise.

1. *Then $U_n''(r)$ may be expressed as the sum of "disjoint" random variables V_k for $k \in \mathbb{N}$. In other words, Ω can be partitioned into measurable subsets E_k such that $V_k = U_n''(r)$ on E_k and $1_\mathscr{G}$ otherwise. Moreover, the V_k may be chosen such that $V_k^*(t) \leqslant k \cdot \ell(t(k-1)!/r^{k-1})$.*
2. *Given the assumptions, for all $p \in (0,\infty)$,*
$$\mathbb{E}_\mu[U_n''(r)^p]^{1/p} \leqslant 2e^{2^p r/p}\mathbb{E}_\mu[\ell_X^p]^{1/p}.$$

With the above results in hand, we can now show the above theorem.

Proof of Theorem 3.6. Compute using the triangle inequality (2.3) and Remark 2.6:

$$d_\mathscr{G}(1_\mathscr{G}, X_k) \leqslant d_\mathscr{G}(1_\mathscr{G}, S_{k-1}) + d_\mathscr{G}(1_\mathscr{G}, S_k) \leqslant 2U_n.$$

Hence $M_n \leqslant 2U_n$. Now compute for $p \geqslant p_0$, using Propositions 3.1 and 3.7:

$$\mathbb{E}_\mu[U_n^p]^{1/p} \geqslant \frac{1}{2}\mathbb{E}_\mu[M_n^p]^{1/p} \geqslant 2^{-1-p_0^{-1}}\mathbb{E}_\mu[\ell_X^p]^{1/p},$$
$$\mathbb{E}_\mu[U_n^p]^{1/p} \geqslant (e^{-p}/8)^{1/p}U_n^*(e^{-p}/8) \geqslant 8^{-p_0^{-1}}e^{-1}U_n^*(e^{-p}/4).$$

Hence there exists a constant $0 < c_1 = c_1(p_0)$ such that:

$$\mathbb{E}_\mu[U_n^p] \geqslant c_1^{-1}(U_n^*(e^{-p}/4) + \mathbb{E}_\mu[\ell_X^p]^{1/p}).$$

This yields one inequality; another one is obtained using Proposition 3.7 as follows:

$$\mathbb{P}_\mu\left(U_n \neq U_n'(e^{-p}/8)\right) \leqslant \mathbb{P}_\mu\left(M_n > \ell_X(e^{-p}/8)\right) \leqslant \mathbb{P}_\mu\left(M_n > M_n^*(e^{-p}/8)\right) \leqslant e^{-p}/8.$$

Now if $\mathbb{P}_\mu\left(U_n'(e^{-p}/8) > y\right) > \eta$ for some $\eta \in [\frac{e^{-p}}{8}, 1]$, then by the reverse triangle inequality,

$$\begin{aligned}\mathbb{P}_\mu\left(U_n > y\right) &\geqslant \mathbb{P}_\mu\left(U_n > y,\ U_n = U_n'(e^{-p}/8)\right)\\ &\geqslant \mathbb{P}_\mu\left(U_n'(e^{-p}/8) > y\right) - \mathbb{P}_\mu\left(U_n \neq U_n'(e^{-p}/8)\right) > \eta - \frac{e^{-p}}{8}.\end{aligned}$$

Hence by definition and the above calculations,

$$U_n'(e^{-p}/8)^*(\eta) \leqslant U_n^*(\eta - e^{-p}/8). \tag{3.11}$$

Applying this with $\eta = e^{-p}/4$,

$$U_n'(e^{-p}/8)^*(e^{-p}/4) \leqslant U_n^*(e^{-p}/8) \leqslant e8^{1/p}\mathbb{E}_\mu[U_n^p]^{1/p} \leqslant e8^{1/p_0}\mathbb{E}_\mu[U_n^p]^{1/p}.$$

Hence as above, there exists a constant $0 < c_2 = c_2(p_0)$ such that:

$$\mathbb{E}_\mu[U_n^p]^{1/p} \geqslant c_2^{-1}(U_n'(e^{-p}/8)^*(e^{-p}/4) + \mathbb{E}_\mu[\ell_X^p]^{1/p}).$$

This proves the second of the four claimed inequalities. The remaining arguments can now be shown by suitably adapting the proof of [6, Theorem 3]. ☐

Finally, we use Theorem 3.6 to prove our remaining main result.

Proof of Theorem B. Using Proposition 2.4, let $\mathscr{G}'$ denote the smallest metric monoid containing $\mathscr{G}$. Thus the X_k are a sequence of independent $\mathscr{G}'$-valued random variables, and we may assume henceforth that $\mathscr{G} = \mathscr{G}'$. Compute using Proposition 3.7, and the fact that X^* and X have the same law for the real-valued random variable $X = M_n$:

$$\begin{aligned}\mathbb{E}_\mu[\ell_X^q] &= \int_0^{1/2} \ell_X(2t)^q \cdot 2dt \leqslant 2\int_0^{1/2} M_n^*(t)^q \; dt \leqslant 2\int_0^1 M_n^*(t)^q \; dt = 2\mathbb{E}_\mu[(M_n^*)^q]\\ &= 2\mathbb{E}_\mu[M_n^q].\end{aligned}$$

Using this computation, as well as Lemma 3.8 and Theorem 3.6 for $\mathscr{G}'$, we compute:

$$\begin{aligned}&\mathbb{E}_\mu[U_n^q]^{1/q}\\ &\leqslant c_1'(\mathbb{E}_\mu[\ell_X^q]^{1/q} + U_n^*(e^{-q}/4))\\ &\leqslant c_1' \cdot 2^{1/q}\mathbb{E}_\mu[M_n^q]^{1/q} + c_1'c_1\frac{\log(4e^q)}{\max(\log(4e^p), \log\log(16e^q))}(U_n^*(e^{-p}/4) + M_n^*(e^{-q}/8))\\ &\leqslant c_1' \cdot 2^{1/q}\mathbb{E}_\mu[M_n^q]^{1/q} + c_1'c_1\frac{\log(4e^q)}{\max(\log(4e^p), \log(\epsilon+q))}(c_2\mathbb{E}_\mu[U_n^p]^{1/p} + M_n^*(e^{-q}/8))\end{aligned}$$

since $\epsilon \in (-q, \log(16)]$. There are now two cases: first if $e^p \geqslant \epsilon + q$, then

$$\frac{\log(4e^q)}{\max(\log(4e^p), \log(\epsilon+q))} \leqslant \frac{q+\log(4)}{p+\log(4)} \leqslant \frac{q}{p} = \frac{q}{\max(p, \log(\epsilon+q))}.$$

On the other hand, if $e^p < \epsilon + q$ then set $C := 1 + \frac{\log(4)}{p_0}$ and note that $Cq \geqslant q + \log(4)$. Therefore,

$$\frac{\log(4e^q)}{\max(\log(4e^p), \log(\epsilon+q))} \leqslant \frac{q+\log(4)}{\log(\epsilon+q)} \leqslant \frac{Cq}{\log(\epsilon+q)} = C\frac{q}{\max(p, \log(\epsilon+q))}.$$

Using the above analysis now yields:

$$\begin{aligned}&\mathbb{E}_\mu[U_n^q]^{1/q}\\ &\leqslant c_1' \cdot 2^{1/q}\mathbb{E}_\mu[M_n^q]^{1/q} + c_1'c_1\left(1 + \frac{\log(4)}{p_0}\right)\frac{q}{\max(p, \log(\epsilon+q))}(c_2\mathbb{E}_\mu[U_n^p]^{1/p} + M_n^*(e^{-q}/8)).\end{aligned}$$

Setting $c := c_1' \max(2^{1/q}, c_1(1+\log(4)/p_0), c_1c_2(1+\log(4)/p_0))$, we obtain the first inequality claimed in the statement of the theorem.

To show the second inequality, we first verify that if $\epsilon \geqslant \min(1, e - p_0)$, then the function $f(x) := x/\log(\epsilon + x)$ is strictly increasing on (p_0, ∞). Now compute:

$$\frac{q}{\max(p, \log(\epsilon+q))} = \min\left(\frac{q}{p}, \frac{q}{\log(\epsilon+q)}\right) \geqslant \min\left(1, \frac{q}{\log(\epsilon+q)}\right) \geqslant \min\left(1, \frac{p_0}{\log(\epsilon+p_0)}\right).$$

Next, use Proposition 3.1 to show: $M_n^*(e^{-q}/8) \leqslant \mathbb{E}_\mu[M_n^q]^{1/q}(8e^q)^{1/q} \leqslant 8^{1/p_0} e \mathbb{E}_\mu[M_n^q]^{1/q}$. Using the previous two facts, we now complete the proof of the second inequality by beginning with the first inequality:

$$\begin{aligned}
&\mathbb{E}_\mu[U_n^q]^{1/q} \\
&\leqslant c\frac{q}{\max(p, \log(\epsilon+q))}(\mathbb{E}_\mu[U_n^p]^{1/p} + M_n^*(e^{-q}/8)) + c\mathbb{E}_\mu[M_n^q]^{1/q} \\
&\leqslant c\frac{q}{\max(p, \log(\epsilon+q))}(\mathbb{E}_\mu[U_n^p]^{1/p} + 8^{1/p_0} e\mathbb{E}_\mu[M_n^q]^{1/q}) + c \cdot 1 \cdot \mathbb{E}_\mu[M_n^q]^{1/q} \\
&\leqslant c\frac{q}{\max(p, \log(\epsilon+q))}\left(\mathbb{E}_\mu[U_n^p]^{1/p} + 8^{1/p_0} e\mathbb{E}_\mu[M_n^q]^{1/q} + \max(1, \frac{\log(\epsilon+p_0)}{p_0})\mathbb{E}_\mu[M_n^q]^{1/q}\right).
\end{aligned}$$

The second inequality in the theorem now follows. □

Acknowledgments. We thank David Montague and Doug Sparks for providing detailed feedback on an early draft of the paper, which improved the exposition.

References

[1] I. Csiszár, *On infinite products of random elements and infinite convolutions of probability distributions on locally compact groups*, Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete (Probability theory and related fields) **5** (1966), no. 4, 279–295.

[2] R.M. Dudley, *Real Analysis and Probability*, Cambridge studies in advanced mathematics **74**, Cambridge University Press, Cambridge-New York, 2002.

[3] A.R. Galmarino, *The equivalence theorem for compositions of independent random elements on locally compact groups and homogeneous spaces*, Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete (Probability theory and related fields) **7** (1967), no. 1, 29–42.

[4] E. Giné and J. Zinn, *Central Limit Theorems and Weak Laws of Large Numbers in Certain Banach Spaces*, Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete (Probability theory and related fields) **62** (1983), no. 3, 323–354.

[5] P. Hitczenko, *On a domination of sums of random variables by sums of conditionally independent ones*, Annals of Probability **22** (1994), no. 1, 453–468.

[6] P. Hitczenko and S.J. Montgomery-Smith, *Measuring the magnitude of sums of independent random variables*, Annals of Probability **29** (2001), no. 1, 447–466.

[7] J. Hoffmann-Jørgensen, *Sums of independent Banach space valued random variables*, Studia Mathematica **52** (1974), no. 2, 159–186.

[8] J. Hoffmann-Jørgensen and G. Pisier, *The law of large numbers and the Central Limit Theorem in Banach spaces*, Annals of Probability **4** (1976), 587–599.

[9] K. Itô and M. Nisio, *On the convergence of sums of independent Banach space valued random variables*, Osaka Journal of Mathematics **5** (1968), 35–48.

[10] W.B. Johnson, G. Schechtman, and J. Zinn, *Best constants in moment inequalities for linear combinations of independent and exchangeable random variables*, Annals of Probability **13** (1985), no. 1, 234–253.

[11] A. Khare and B. Rajaratnam, *Differential calculus on the space of countable labelled graphs*, preprint (arXiv:1410.6214), 2014.

[12] A. Khare and B. Rajaratnam, *Integration and measures on the space of countable labelled graphs*, preprint (arXiv:1506.01439), 2015.

[13] A. Khare and B. Rajaratnam, *The Hoffmann-Jørgensen inequality in metric semigroups*, Annals of Probability, to appear, 2016.

[14] A. Khare and B. Rajaratnam, *The Khinchin–Kahane inequality and Banach space embeddings for metric groups*, Technical Report, Stanford University, 2016.

[15] V.L. Klee Jr., *Invariant metrics in groups (solution of a problem of Banach)*, Proceedings of the American Mathematical Society **3** (1952), no. 3, 484–487.

[16] P. Kranz and M.-Guo, *A metrizable cancellative semigroup without translation invariant metric*, Proceedings of the American Mathematical Society **66** (1977), 17–19.

[17] S. Kwapień and W.A. Woyczyński, *Random series and stochastic integrals. Single and multiple*, Birkhäuser, Boston, 1992.

[18] M. Ledoux and M. Talagrand, *Probability in Banach Spaces (Isoperimetry and Processes)*, Ergebnisse der Mathematik und ihrer Grenzgebiete, Springer-Verlag, Berlin-New York, 1991.

[19] L. Lovász, *Large Networks and Graph Limits*, volume 60 of *Colloquium Publications*, American Mathematical Society, Providence, 2012.

[20] A.A. Mogul'skii, *On the law of the iterated logarithm in Chung's form for functional spaces*, Theory of Probability and its Applications **24** (1979), no. 2, 405–413.

[21] A. Mukherjea and T.C. Sun, *Convergence of products of independent random variables with values in a discrete semigroup*, Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete (Probability theory and related fields) **46** (1978), no. 2, 227–236.

[22] H.P. Rosenthal, *On the subspaces of L^p $(p > 2)$ spanned by sequences of independent random variables*, Israel Journal of Mathematics **8** (1970), no. 3, 273–303.

[23] A. Tortrat, *Lois de probabilite sur un espace topologique completement régulier et produits infinis à termes indépendant dans un groupe topologique*, Annales de l'Institut Henri Poincaré (B): Probabilités et Statistique **1** (1964/65), 217–237.

Department of Mathematics and Department of Statistics, Stanford University, Stanford, CA 94305, USA

E-mail address: `A.K.: khare@stanford.edu; B.R.: brajaratnam01@gmail.com`